

Modelling the long wave-phenomena

Metz, Rainer; Stier, Winfried

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Metz, R., & Stier, W. (1992). Modelling the long wave-phenomena. *Historical Social Research*, 17(3), 43-62. <https://doi.org/10.12759/hsr.17.1992.3.43-62>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see:
<https://creativecommons.org/licenses/by/4.0>

Modelling the Long Wave-Phenomena

*Rainer Metz / Winfried Stier**

Abstract: In this paper, by »modelling« we mean the identification and estimation of time series models and not the development and evaluation of economic-theoretical models. Whereas the latter mentioned approach aims at analysing the probable causes of wave-phenomena, the time-series approach is a purely empirical one. If it should prove to be possible to identify time-series models, we can hope that they will possess predictive power and we can further hope that this will help us in finding dependencies between different wave-series. This would make it possible to conduct multivariate analyses of wave-phenomena which have to the best of our knowledge not been performed yet. However, in this paper we restrict ourselves to univariate modelling for two reasons: in the first place, we think that sufficient experience must be accumulated in univariate modelling before multivariate modelling can be done properly, and secondly, even univariate modelling of long wave-phenomena by means of modern time series analysis is a topic not discussed up to now (as far as we know).

1. Statistical Identification of Long Waves - the Filter Problem

We agree with Menshikov/Klimenko [2] when they say: »The statistical identification of long waves is not a trivial problem and needs refined mathematical methods«. Indeed, time series-modelling of wave-phenomena seems to be a priori only successful if it is possible to isolate wave-movements, that is, they should not be buried in background-noise. It is not hard to show that the traditional tools of time series analysis are in

* Address all communications to Rainer Metz, Zentralarchiv für Empirische Sozialforschung - Abt. ZHSF, Universität zu Köln, Bachemerstr. 40, D-5000 Köln 41.

general not suitable for this purpose. The use of moving averages, differences and least-squares estimates - to mention the most often used approaches - has serious drawbacks in this context. For instance, differences (of first or higher order) have the property of eliminating or at least damping low-frequency oscillations - the very oscillations we are interested in. Besides that they tend to change the periodicities of these oscillations, more exactly: they shorten the periods of these oscillations since the amplitude-functions of difference-filters have a positive slope for all frequencies.

Moving averages and least-squares have similar and/or other drawbacks which are well-known and therefore need not be discussed here.

The isolation of long waves requires the design of band-pass filters. Such filters have amplitude-functions showing the value one in the pass-band and zero in the stopbands. Besides that two (small) transition-bands have to be specified (this is necessary to avoid undesirable side-effects like the Gibb's phenomenon).

Now, the design of digital filters is a relatively complicated matter which is impossible to discuss here fully and in detail. The interested reader is referred to Stier [3, 4, 5]. Only some basic facts can be presented here. Generally speaking, there are two different types of filters, FIR (= Finite Impulse Response)- and IIR (= Infinite Impulse Response)-Filters. FIR-Filters have a finite and IIR-Filters an infinite impulse response, which is the filter-output if the filter-input is the discrete impulse-function:

$$\delta_t = \begin{cases} 0, & t \neq 0 \\ 1, & t = 0 \end{cases}$$

The filter equation for linear filters generally is:

$$y_t = \sum_{k=0}^M a_k y_{t-k} + \sum_{r=0}^N b_r x_{t-r}, \quad a_0 = 1$$

For FIR-filters $a_1=a_2=\dots=a_M=0$ holds. Obviously, moving averages and differences are special FIR-filters. Filters are completely determined when the coefficients a_i and b_i are known. The properties of a filter in the frequency-domain are given by its transfer-function

$$T(\lambda) = \frac{\sum_{r=0}^N b_r e^{-i\lambda r}}{\sum_{k=0}^M a_k e^{-i\lambda k}}, \quad -\pi \leq \lambda \leq \pi.$$

$T(\lambda)$ is in general a complex-valued function which can be written in the form

$$T(\lambda) = |T(\lambda)| e^{i\varphi(\lambda)}$$

where $|T(\lambda)|$ is the amplitude- and, $\varphi(\lambda)$ the phase-function of the filter. The amplitude-function shows which frequencies (or frequency-bands) are dampened (or eliminated) or amplified and the phase-function gives informations about how many time units the filter-output lags behind the filter-input. Filter-design consists essentially in determining the filter-coefficients in such a manner that a desired amplitude- and/or phase-function results. There are various design-approaches which differ in complexity and efficiency.

FIR-filter and IIR-filter differ mainly in two - practically important - aspects. It is possible to design IIR-filters with a very »sharp« amplitude-characteristic even with a low filter-order (i.e. small N and M). Equivalent FIR-filters are only possible with a very high filter-order (i.e. great N), which is usually so large (e.g. $N=100$) that these filters are of no use in analysing time series of only moderate length.

Therefore, for statistical identification of long waves - which requires the design of sharp band-passes - FIR-filters cannot be used. The advantage of FIR-filters (compared to IIR-filters) is their linear phase-characteristic, which includes as a special case zero-phase between filterin- and filter-output (if the filter is implemented symmetrically). IIR-filters have always a non-linear phase-function and therefore never exact zero-phase. Of course, in the case of identifying long waves, zero phase between filter-input and filter-output would be highly desirable, since then maxima, minima and turning-points of waves are not distorted on the time scale. But unfortunately, the (ideal) requirements of an exact amplitude-characteristic and an exact zero-phase exclude each other. We cannot have both - at least as long as the filtering is done in the time-domain.

Fortunately, there is a third possibility of filter-design. Both FIR-filters and IIR-filters can be characterized by the fact that the filter is designed in the frequency-domain, but the filtering itself is done in the time-domain. However, it is possible, both to design *and* filter in the frequency-domain:

Let $x_t, t=0, 1, \dots, T-1$ be the given time series. Then the Discrete Fourier-Transform (DFT)

$$X(e^{i\lambda}) := \sum_{t=0}^{T-1} x_t e^{-i\lambda t}, \quad \lambda \in [0, 2\pi]$$

represents the complete frequency-content of the given signal x_t . The DFT is defined for a continuum of frequencies. Therefore it seems to be impossible to evaluate $X(e^{i\lambda})$ directly. Furthermore, it seems to be impossible to restore the original signal x_t by an inverse transform. Fortunately, it can be shown that it is sufficient to know $X(e^{i\lambda})$ only for T frequency-points, that is

$$X(e^{i\lambda}) = X_k = \sum_{t=0}^{T-1} x_t e^{-i(2\pi/T)tk}, \quad k = 0, 1, \dots, T-1.$$

in order to restore the signal x_t completely by the so-called »Inverse Discrete Fourier-Transform« (IDFT). Therefore we can write:

$$x_t = \sum_{k=0}^{T-1} X(k) e^{i(2\pi/T)kt}, \quad t = 0, 1, \dots, T-1.$$

This means that a one-to-one-relationship exists between the signal and its DFT, which implies that no information is lost by transforming the signal x_t . Practically, a higher sampling rate of the frequency-intervall $[0, 2\pi]$ is to be recommended, especially for relatively small T . This can be done by defining a new signal of length $L > T$

$$\tilde{x}_t := \begin{cases} x_t, & t = 0, 1, \dots, T-1 \\ 0, & T \leq t \leq L-1 \end{cases}$$

whose DFT of length L is

$$\tilde{x}(e^{i\lambda}) = \tilde{X}_k = \sum_{t=0}^{L-1} \tilde{x}_t e^{-i(2\pi/L)tk}, \quad k = 0, 1, \dots, L-1$$

which is defined for the frequencies $\lambda_k = (2\pi/L)k, k=0, 1, \dots, L-1$. Since $\tilde{x}_t=0$ for $t > T$ we can write

$$\bar{x}_k = \sum_{t=0}^{T-1} x_t e^{-i(2\pi/L)tk} , \quad k = 0, 1, \dots, L - 1.$$

The filtering procedure seems to be very simple using this approach: specify the amplitude-function of the defined filter in the frequency-domain, multiply the DFT of the signal by it and apply the IDFT to get the filtered series in the time domain. By this way we realize both ideal amplitude-functions *and* zero-phase between filter-input and filter-output. Unfortunately, things are not so simple as they seem to be: The »first« and »last« values of the filtered series are more or less distorted. There is a special »boundary-value-problem«. It can be shown that this problem is due to a »filter-error« which is the result of having only *finite* series in practice and *not infinite* ones as in theory. A detailed analysis shows that this error is the product of two components. The first one is the supremum of the spectrum of the signal x , and the second one depends on the filter-weights (which result via IDFT from the amplitude-function). By a special procedure, this filter-error can be made smaller iteratively. Though it is not possible to get completely distortion-free values at the edges of a series, the smallness of distortion finally reached allows boundary-values whose filter-error can be neglected. For details see Stier [3]. Thus, this filtering approach realizes *both* exact amplitudes and zero-phase. It is to be noted that the filtered values at the right edge of a series do not remain unaltered when the series is updated. Indeed, it can be shown easily that it is not possible to have exact amplitude, zero-phase *and* absolute stability of the filtered series. In contrast, IIR-filters show exact amplitude, absolute stability, but no zero-phase. At most we can have only two of the three desirable properties of filters (exact amplitude, zero-phase, absolute stability) *simultaneously*.

3. The Modelling Approach

The approach used here is the well-known Box-Jenkins-approach. This model building philosophy tries to model a given time series by a so-called ARIMA(p,d,q)-process:

$$a(B)(1-B)^d X_t = b(B)u_t$$

where

$$a(B) = 1 - a_1 B - a_2 B^2 - \dots - a_p B^p$$

$$b(B) = 1 + b_1 B + b_2 B^2 + \dots + b_q B^q.$$

B denotes the backward shift-operator with the property

$$B^k X_t = X_{t-k}$$

and u_t is a white-noise process. (The primary reference is Box and Jenkins [1]).

In this approach it is assumed that a given series is a realization of a specific non-stationary stochastic process which can be transformed to stationarity by simple differencing. d denotes the degree of differencing ($d=1, 2, \dots$) and $a(B)$ resp. $b(B)$ the autoregressive resp. the moving average-operator. Defining

$$Y_t = (1 - B)^d X_t$$

we can write

$$a(B) Y_t = b(B) u_t$$

or

$$\sum_{j=0}^p a_j Y_{t-j} = \sum_{k=0}^q b_k u_{t-k}, \quad a_0 = 1$$

which means that the differenced series can be represented by a stationary mixed autoregressive-moving average (ARMA-)-model. A strategy for constructing ARIMA-models can be based on a three-step iterative cycle of

- a) model identification
- b) model estimation
- c) diagnostic checks on model adequacy.

It would lead us much too far away to discuss even the basics of these steps here. Suffice it to say that in step a) the estimated autocorrelation- and partial autocorrelation-function are important guides in search of the »correct« model and that in step c) residual analysis is most important (different tests for white-noise etc.).

Now, the original approach of Box/Jenkins as sketched above very briefly is not suitable for modelling long-wave phenomena: the required differencing to achieve stationarity eliminates (or at least dampens and shifts) long cycles as mentioned above already. For instance, the »best« model for our series is

$$(1 - B) x_t = (1 + 0.47B + 0.15B^2 + 0.24B^3 + 0.19B^4) u_t$$

or

$$X_t = X_{t-1} + u_t + 0.47u_{t-1} + 0.15u_{t-2} + 0.24u_{t-3} + 0.19u_{t-4}$$

This is an IMA(1,4)-process, which shows a good fit to the given data. However, this process is not able to show any cyclic behaviour, since for this purpose the autoregressive part must be at least of order two. This result is not surprising if we look at the series of the first differences which is shown in figure 1. Obviously, high-frequency-movements dominate the series. Using higher differences would accentuate this even more (see figure 2 for a second order difference). Therefore, only highpass- and/or bandpass-filtered series promise interesting results.

4. Empirical results

By filtering our series »Gross Fixed Capital Formation« with a high-pass whose pass-band is the frequency-interval $[0.0125, 0.5]$ - which means that all oscillations with periods of 80 to two years are left untouched - we get figure 3 (compared to first differences which are also plotted). Obviously, the series is trend-free now. But this does not imply that we are allowed to consider it as a realization of a stationary ARMA-process. Stationarity means also homoscedasticity, i.e. a constant variance. Obviously, the detrended series shows a strongly increasing variance. In cases like this, variance-stabilizing transforms (such as the Box-Cox transforms) are necessary before modelling can take place. We did not use any of these transforms, but put the detrended values in relation to the trend of the original series, a transform which is easy to interpret. This trend was determined by a low-pass filter whose pass-band was the frequency-interval $[0, 0.0125]$, which implies that all oscillations with a period longer than 80 years are defined to be »trend«. Figure 4 shows the trend and the bandpass-filtered series together with the original values.

The pass-band of the band-pass is the frequency-interval $[0.0125, 0.05]$. This means that all oscillations with a period of 80 to 20 years are preserved by the filter. These are the long waves we are trying to model.

As the graphs show, the variance could be stabilized, although there is still some doubt if homoscedasticity can be assumed, especially in the (relative) trendfree series (which is not shown here).

In fact, it proved to be difficult to find an ARMA-model of low order fitting satisfactorily the (relative) trendfree series. The best fit showed an autoregressive model of order $p = 18$. This result is not very useful, especially considering later multivariate analysis. But we admit that we did not spend much effort to find a simpler model, since our target is the modeling of the proper long wave, i.e. the band-pass filtered series.

Figure 5 shows the estimated autocorrelation function of the band-pass filtered series (the »long wave«). This function clearly shows a cyclical pattern, indicating the presence of a strong cycle in the series. The estimated partial autocorrelation function showed significant values only for the first two lags which suggests an autoregressive part of order two. The best model proved to be the ARMA(2,1)-process:

$$\begin{array}{rcc}
 X_t - 1.889x_{t-1} + 0.943X_{t-2} = u_t + 0.453u_{t-1} \\
 (0.032) \quad (0.032) \quad (0.084) \\
 (t=59.0) \quad (t=29.5) \quad (t=5.4)
 \end{array}$$

The estimated standard errors (shown in parantheses) are very small and the t-values large. The fit of the model is excellent, the long wave-series and the series resulting from the model are hardly discernible (see figure 6).

The adequacy of the ARMA(2,1)-model can be tested further by analyzing the autocorrelation function and the spectrum of the process and by comparing the results with the estimated spectrum of the long wave.

If we assume that the »true« model of the long wave is given by an ARMA(2,1)-process with the coefficients $a_1 = -1.889$, $a_2 = 0.943$, $b_1 = 0.453$, then we can derive the autocorrelation function of this process. Generally, for an ARMA(p,q)-process the autocorrelations R_τ obey

$$R_\tau = - \sum_{j=1}^p a_j R_{\tau-j}, \quad \tau > q$$

which reduces in our case to the second-order difference equation:

$$R_\tau - 1.889 R_{\tau-1} + 0.943 R_{\tau-2} = 0, \quad \tau > 1$$

The characteristic equation of this homogenous difference equation is

$$\lambda^2 - 1.889 \lambda + 0.943 = 0,$$

which has the roots

$$\lambda_{1/2} = -0.9445 \pm 0.2256541 \cdot i.$$

The solution of this equation is given by

$$R_t = A r^t \cos(\theta t - \epsilon)$$

with

$$r = |\lambda_{1/2}| = 0.971089$$

$$\theta = \arctg \left(- \frac{\sqrt{4 a_2 - a_1^2}}{a_1} \right)$$

$$= 0.2345177.$$

The periodicity of R_t is $P = 2\pi / \theta \sim 26.8$ years. The constants A and ϵ depend on the initial values of R_t , namely $R_0=1$ and R_1 . Since R_1 is not known, they cannot be determined, which however does not matter here (if we would replace R_1 by the estimate $R_1=0.966$ we would get $A \sim 1.0436$ and $\epsilon \sim 0.0923$ and a high similarity between the »model-correlations« R_t and the estimates R_t could be observed).

The spectrum of an ARMA(2,1)-process is:

$$S(\lambda) = \frac{\sigma^2}{2\pi} \frac{1 + 2b_1 \cos \lambda + b_1^2}{1 + a_1^2 + a_2^2 + 2(a_1 + a_1 + a_2) \cos \lambda + 2a_2 \cos 2\lambda}$$

and

$$S(\lambda) = \frac{\sigma^2}{2\pi} \frac{1.205 + 0.906 \cos \lambda}{5.46 - 7.34 \cos \lambda + 1.886 \cos 2\lambda} \quad |\lambda| \leq \pi$$

is its »peak-frequency«.

$$\lambda_p = 0.2327 \quad \frac{1.205 + 0.906 \cos \lambda}{5.46 - 7.34 \cos \lambda + 1.886 \cos 2\lambda} \quad |\lambda| \leq \pi$$

which indicates a cycle of 26 years. Figure 7 shows the *estimated* spectrum of the long wave-series (in fact the periodogram is shown, obtained after tapering by using a cosine-bell. Since the difference between spectrum and periodogram can be neglected for series of the long wave-type, we used for the sake of simplicity the periodogram). The spectrum shows a clear peak at the frequency $\lambda_p=0.2454$ which corresponds to a cycle-length of about 26 years. Thus the cycle-length of the long wave-series in the two approaches

is practically the same. Since the time-series model and the estimation of the spectrum (which is a non-parametric approach) lead to practically identical results concerning the periodicity of the wave, the result of the modelling-procedure - the ARMA(2,1)-process - can be considered to be a valid one.

In addition, we tried to get informations about the temporal stability of the identified ARMA-model. For this purpose, we divided the long wave-series into two parts. The first part ends at 1912 (length 83). This time-point was chosen - maybe somewhat arbitrarily - because it was felt that the wave-form changes from there on.

For the first part we identified the ARMA(2,1)-process

$$\begin{aligned}
 X_t^{(1)} - 1.879 X_{t-1}^{(1)} + 0.952 X_{t-2}^{(1)} &= u_t + 0.589 u_{t-1} \\
 (0.06) \quad (0.056) \quad (0.165) \\
 (t = 31.3) \quad (t = 57.0) \quad (t = 3.6)
 \end{aligned}$$

If we consider this to be the »true« model again, we can derive a »model-autocorrelation« function and a »model-spectrum« as we did for the complete series. Here we get

$$\begin{aligned}
 R_\tau^{(1)} &= A \cdot 0.975705^\tau \cos(0.2732697\tau - \epsilon) , \\
 R_\tau^{(1)} &= A \cdot 0.975705^\tau \cos(0.2732697\tau - \epsilon) ,
 \end{aligned}$$

$$S^{(1)}(\lambda) = \frac{\sigma^2}{2\pi} \frac{1.347 + 1.178 \cos \lambda}{5.44 - 7.34 \cos \lambda + 1.904 \cos 2\lambda}$$

i.e. the periodicity of R_τ is $P=23$ years. The spectrum is

$$\lambda_P^{(1)} = 0.2722$$

$$S^{(1)}(\lambda) = \frac{\sigma^2}{2\pi} \frac{1.347 + 1.178 \cos \lambda}{5.44 - 7.34 \cos \lambda + 1.904 \cos 2\lambda}$$

with

$$\lambda_P^{(1)} = 0.2722$$

$$X_t^{(2)} - 1.9116 X_{t-1}^{(2)} + 0.9060 X_{t-2}^{(2)} = u_t + 0.700 u_{t-1}$$

$$(0.044) \quad (0.042) \quad (0.114)$$

$$(t = 43.5) \quad (t = 22.9) \quad (t = 6.1)$$

$$R_t^{(2)} = A \cdot 0.978767^\tau \cos(0.2063637\tau - \epsilon)$$

$$(P = 30)$$

$$S^{(2)}(\lambda) = \frac{\sigma^2}{2\pi} \frac{1.49 + 1.4 \cos 2\lambda}{5.59 - 7.51 \cos \lambda + 1.92 \cos 2\lambda}$$

$$(\lambda_p^{(2)} = 0.2103479, \text{ i.e. } 30 \text{ years}).$$

The estimated spectrum (periodogram) indicates a cycle of 32 years. Thus, the agreement is still a very good one, but not as perfect as in the first part (a closer look at the periodogram shows that the periodogram-ordinates are practically the same for the 8th and 9th frequency-point (corresponding to 32 resp. 28 years). Therefore, a cycle-length of 30 years could also be justified with good reasons.

Obviously, the cycle-length of 26 years for the complete long wave-series is an average-information. If we take a weighted mean of both cycle-lengths (which is maybe somewhat naive) we get

$$\frac{83 \cdot 23 + 67 \cdot 30}{150} \approx 26,$$

i.e. the cycle-length of the complete series.

One might wonder, if it is possible to detect a cycle by an estimated spectrum (or periodogram), which is contained only two or three times in a series. Under certain circumstances this is possible indeed. For example, let us consider the simulated series $Z_t = 100 \sin(2\pi \cdot 0.03125 t)$ (cycle-length = 32). If we take only 32 values, which means that the cycle is contained in the series only once, then we get a periodogram in which the cycle is indicated by a large (though relatively »broad«) peak. If the series gets longer, this peak gets larger and »narrower«. This result would not be possible, if the signal Z_t were »buried« in noise (say trend and high-frequency components). It is only possible when signals are available »in nuce« (like our Long Wave-Series). To us, it seems to be hopeless trying to analyze and/or model long cycles by using original series or transformed

ones by means of the traditional (and useless in this context) filter-approaches.

Let us finally check the predictive power of our ARMA(2,1)-process. For this purpose our series »Gross fixed capital formation« is updated until 1985 and band-pass filtered with the same filter as before. Then we get 6 new long wave-values. The values forecasted by the two ARMA(2,1)-processes (for the complete series and the second part of it) and the »true« values are found in the following table:

Year	Forecasted Values	Forecasted Values (2nd Part of Series)	»True« Values
1980	-12.4	-12.4	-9.0
1981	-13.0	-13.3	-11.5
1982	-13.0	-13.6	-13.2
1983	-12.2	-13.2	-13.9
1984	-10.9	-12.3	-13.6
1985	-9.0	-10.9	-12.4

Comparing the forecasted values with the »true« ones, it is evident that both model-forecasts show the correct tendency: there is a »turning-point« (where the downswing stops) both in the »true« data and in the forecasting. The latter dates this point one year too early, however. Besides that, the minima are very close together in value. The second forecast is somewhat superior, the values are closer to the »true« ones. This is to be expected of course.

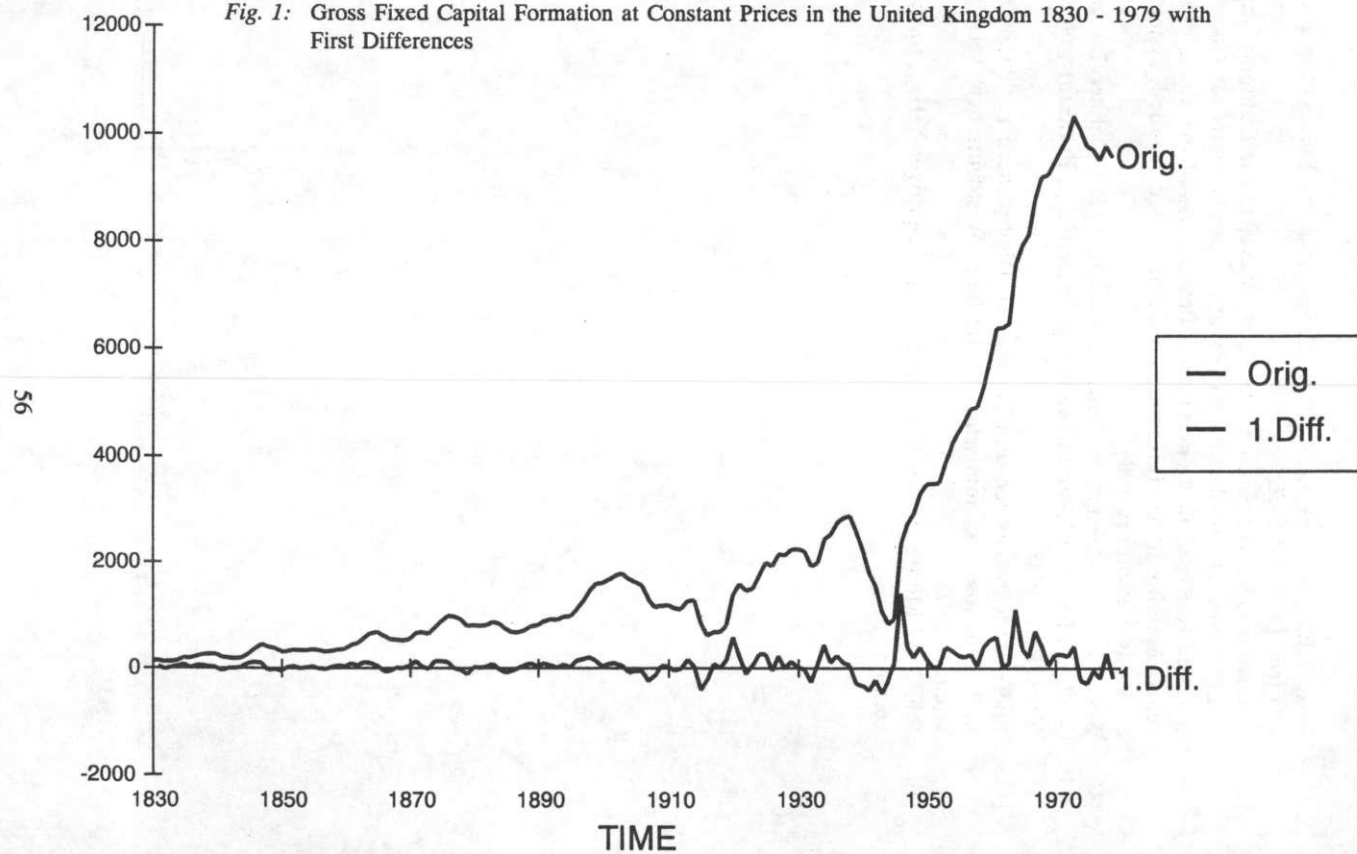
5. Summary

In this paper we have shown how long wave-informations can be extracted from a given time series by using distortion-free filter-techniques. We have further demonstrated how wave-phenomena can be modelled by means of modern time series analysis. Finally, it was shown that the parametric models found this way have predictive power (at least for the series we used). However, we think that much work still needs to be done to gain a more thorough understanding of the problems involved in modelling wave-phenomena. This paper is to be considered as a first step only.

References

- [1] Box, G.E.P.; Jenkins, G.M.: Time Series Analysis, Forecasting and Control, Oakland 1976.
- P] Menshikov, S.; Klimenko, L.: On Long Waves in the Economy, in: Long Waves, Depression And Innovation: Implications for National And Regional Economic Policy, Proceedings of the Siena/Florence Meeting 1983, International Institute for Applied System Analysis, Laxenburg 1985
- [3] Stier, W.: Filter-Design im Frequenzbereich, in: H.J. Gerster: Lange Wellen wirtschaftlicher Entwicklung, Peter Lang, Frankfurt, New York, Paris 1988.
- [4] Stier, W.: Konstruktion und Einsatz von Digitalfiltern zur Analyse und Prognose ökonomischer Zeitreihen, Westdeutscher Verlag, Opladen 1978.
- [5] Stier, W.: Methoden der Zeitreihenanalyse, Springer-Verlag, forthcoming.

Fig. 1: Gross Fixed Capital Formation at Constant Prices in the United Kingdom 1830 - 1979 with First Differences



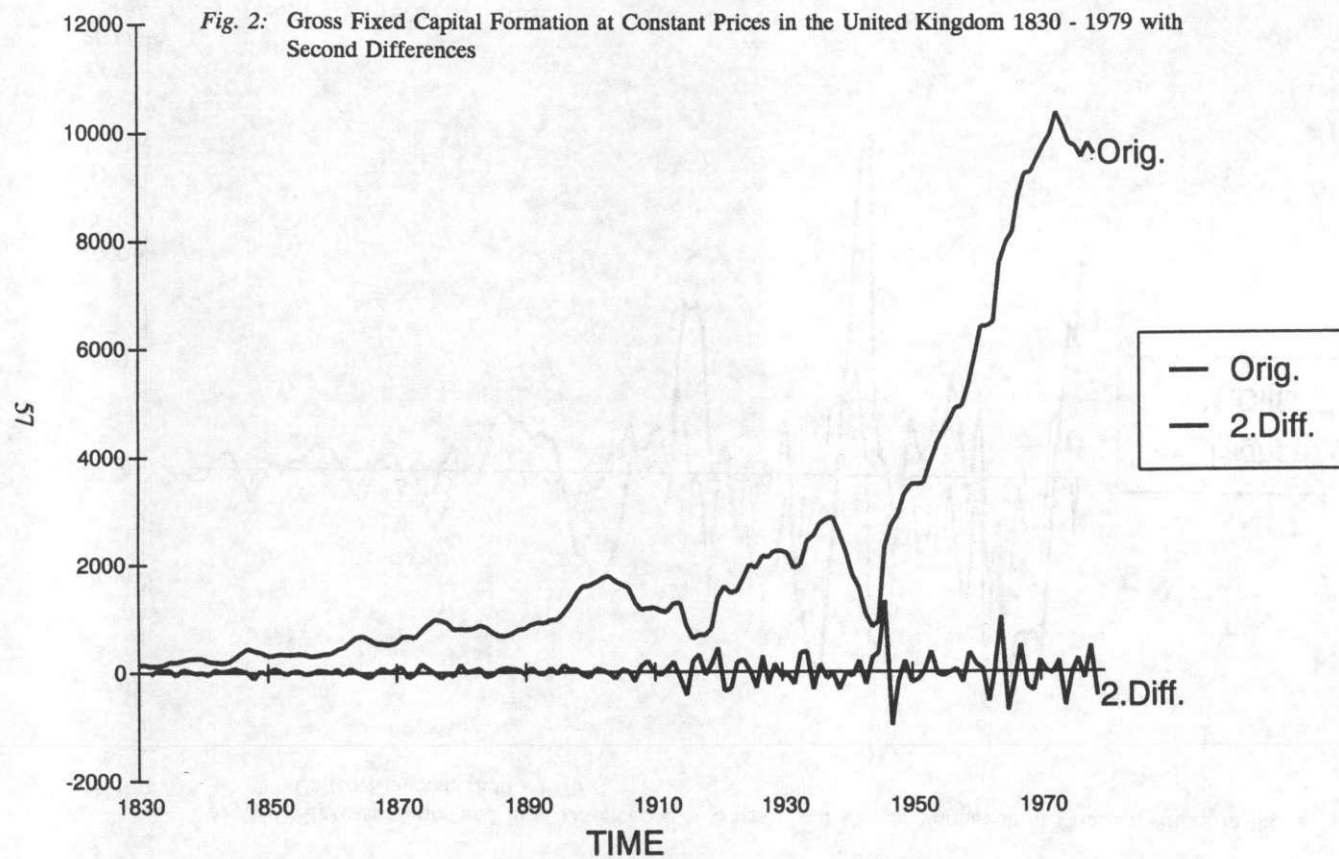


Fig. 3: Highpass Series and First Differences of Gross Fixed Capital Formation at Constant Prices in the United Kingdom 1830 - 1979

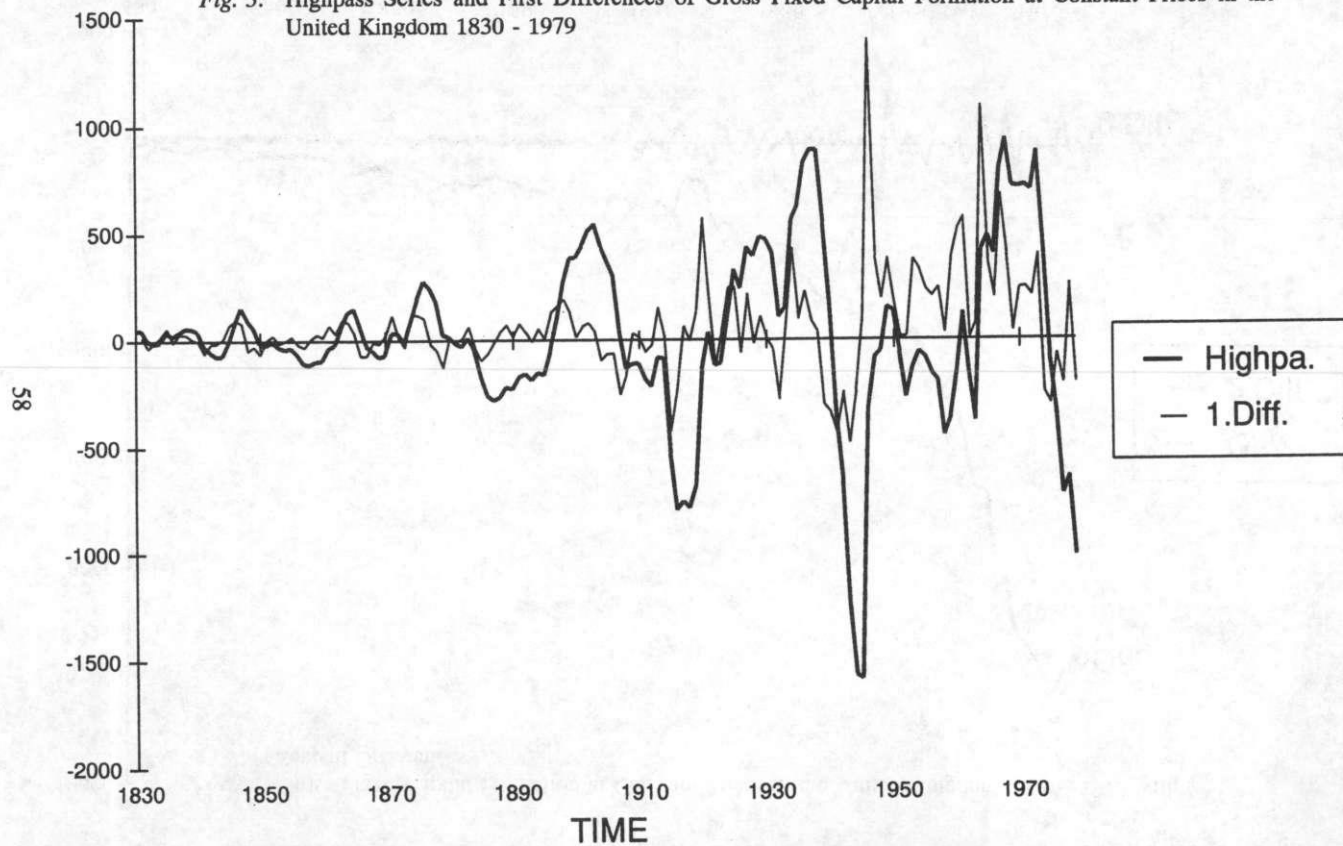
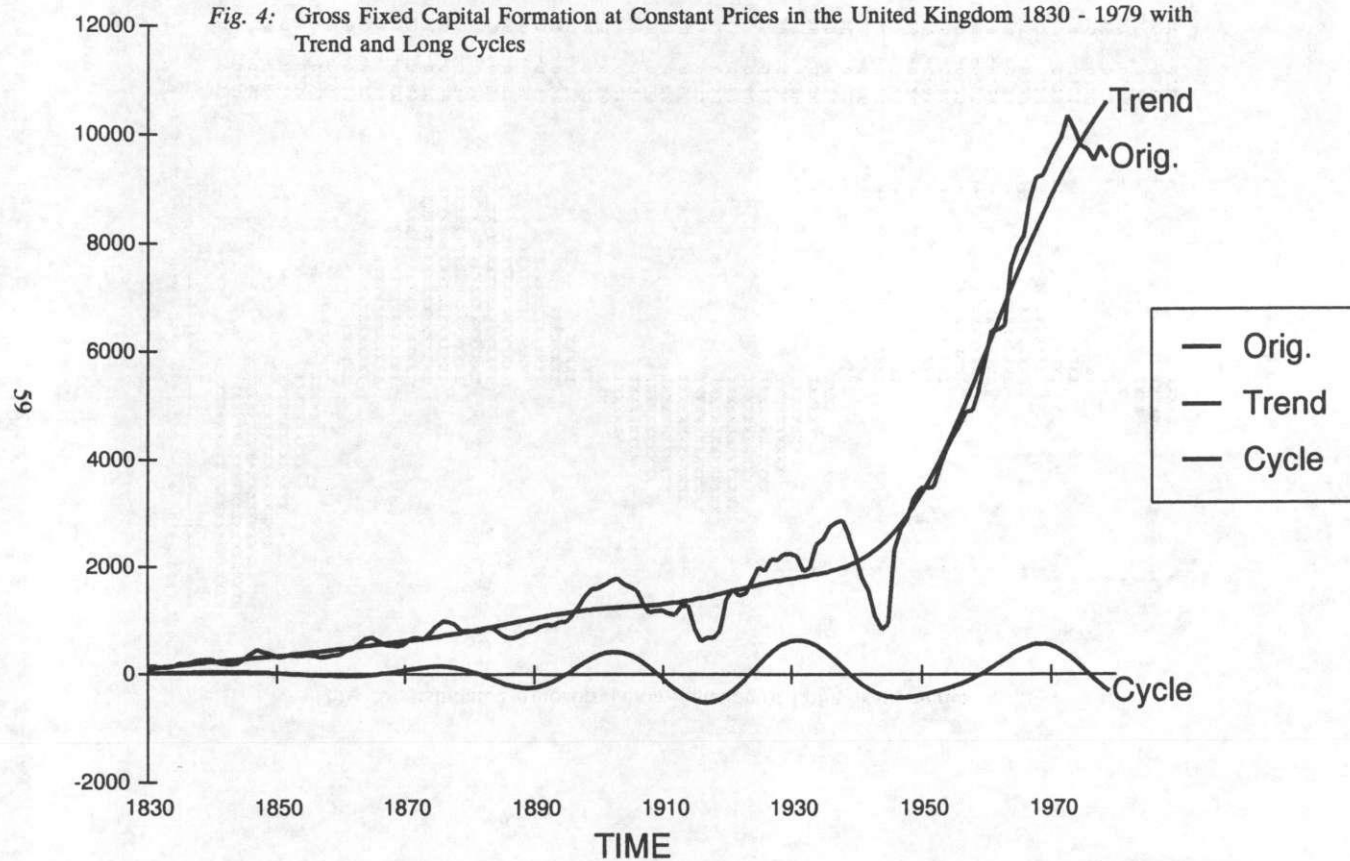
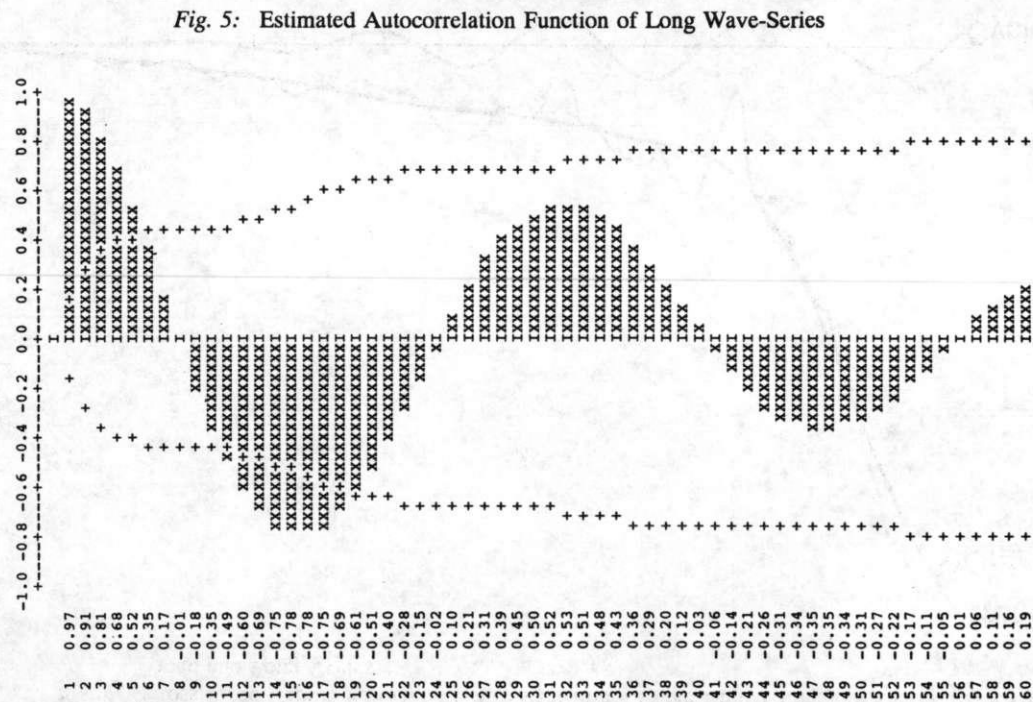


Fig. 4: Gross Fixed Capital Formation at Constant Prices in the United Kingdom 1830 - 1979 with Trend and Long Cycles





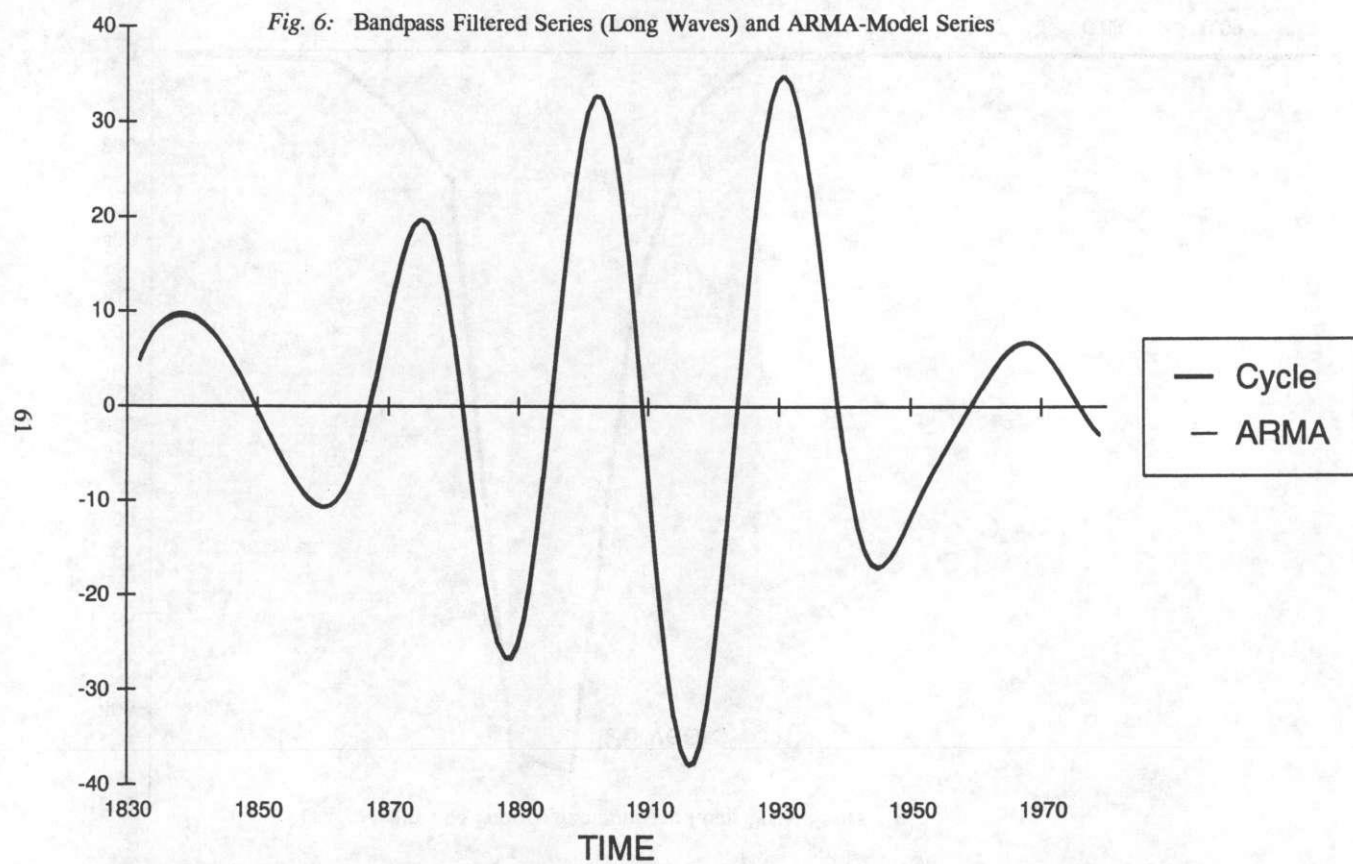


Fig. 7: Estimated Periodogramm of the Long Wave-Series

